AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

METHODS

Sybr Green- and TaqMan-Based Quantitative PCR Approaches Allow Assessment of the Abundance and Relative Distribution of *Frankia* Clusters in Soils

Seifeddine Ben Tekaya,[a] Abirama Sundari Ganesan,[a] Trina Guerra,[a] Jeffrey O. Dawson,[b] Michael R. J. Forstner,[a] Dittmar Hahn[a]

Texas State University, Department of Biology, San Marcos, Texas, USA[a]; University of Illinois at Urbana-Champaign, Department of Natural Resources and Environmental Sciences, Urbana, Illinois, USA[b]

ABSTRACT The nodule-forming actinobacterial genus *Frankia* can generally be divided into 4 taxonomic clusters, with clusters 1, 2, and 3 representing nitrogen-fixing strains of different host infection groups and cluster 4 representing atypical, generally non-nitrogen-fixing strains. Recently, quantitative PCR (qPCR)-based quantification methods have been developed for frankiae of clusters 1 and 3; however, similar approaches for clusters 2 and 4 were missing. We amended a database of partial 23S rRNA gene sequences of *Frankia* strains belonging to clusters 1 and 3 with sequences of frankiae representing clusters 2 and 4. The alignment allowed us to design primers and probes for the specific detection and quantification of these *Frankia* clusters by either Sybr Green- or TaqMan-based qPCR. Analyses of frankiae in different soils, all obtained from the same region in Illinois, USA, provided similar results, independent of the qPCR method applied, with abundance estimates of 10×10^5 to 15×10^5 cells (g soil)$^{-1}$ depending on the soil. Diversity was higher in prairie soils (native, restored, and cultivated), with frankiae of all 4 clusters detected and those of cluster 4 dominating, while diversity in soils under *Alnus glutinosa*, a host plant for cluster 1 frankiae, or *Betula nigra*, a related nonhost plant, was restricted to cluster 1 and 3 frankiae and generally members of subgroup 1b were dominating. These results indicate that vegetation affects the basic composition of frankiae in soils, with higher diversity in prairie soils compared to much more restricted diversity under some host and nonhost trees.

IMPORTANCE Root nodule formation by the actinobacterium *Frankia* is host plant specific and largely, but not exclusively, correlates with assignments of strains to specific clusters within the genus. Due to the lack of adequate detection and quantification tools, studies on *Frankia* have been limited to clusters 1 and 3 and generally excluded clusters 2 and 4. We have developed tools for the detection and quantification of clusters 2 and 4, which can now be used in combination with those developed for clusters 1 and 3 to retrieve information on the ecology of all clusters delineated within the genus *Frankia*. Our initial results indicate that vegetation affects the basic composition of frankiae in soils, with higher diversity in prairie soils compared to much more restricted diversity under some host and nonhost trees.

KEYWORDS alder, birch, qPCR, quantification, saprotrophic, soil

Received 12 October 2016 **Accepted** 9 December 2016

Accepted manuscript posted online 16 December 2016

Citation Ben Tekaya S, Ganesan AS, Guerra T, Dawson JO, Forstner MRJ, Hahn D. 2017. Sybr Green- and TaqMan-based quantitative PCR approaches allow assessment of the abundance and relative distribution of *Frankia* clusters in soils. Appl Environ Microbiol 83:e02833-16. https://doi.org/10.1128/AEM.02833-16.

Editor Claire Vieille, Michigan State University

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Dittmar Hahn, dh49@txstate.edu.

Frankiae are slow-growing actinobacteria that are able to form root nodules with some woody nonleguminous plants (1–3). Root nodule formation is host plant specific and largely, but not exclusively, correlates with assignments of strains to

specific clusters derived from comparative analyses of 16S rRNA gene sequences (4). Cluster 1 represents frankiae that form nodules on plants from the genera *Alnus*, *Morella*, *Myrica*, and *Comptonia* and includes a subgroup infecting the genera *Casuarina* and *Allocasuarina* (5, 6). Members of cluster 2 represent frankiae nodulating *Dryas*, *Purshia*, *Chamaebatia*, *Cercocarpus*, *Ceanothus*, *Datisca*, and *Coriaria*, while members of cluster 3 form nodules on plants that include the genera *Elaeagnus*, *Hippophaë*, *Shepherdia*, *Myrica*, *Morella*, *Gymnostoma*, *Discaria*, *Trevoa*, *Retanilla*, *Kentrothamnus*, and *Colletia* (5, 6). In addition to the typical nitrogen-fixing frankiae, atypical, generally non-nitrogen-fixing and/or nonnodulating frankiae have been identified in cluster 4 (4).

Since the first report of an isolation of *Frankia* from root nodules in 1978 (7), a large number of isolates has been obtained for clusters 1 and 3 (8–11) and a few for cluster 4 (12, 13). An isolate representing cluster 2, however, has been obtained only recently (14). Consequently, most studies on *Frankia* have been performed on those from clusters 1 and 3 and far less on frankiae of clusters 2 and 4. Information on the ecology of cluster 2 and 4 frankiae therefore is quite limited. We have recently developed Sybr Green-based quantitative PCR (qPCR) methods that used *nifH* or 23S rRNA genes as a target to quantify uncultured *Frankia* populations in different soils (15–17). *nifH* as a target only detected the combination of members of clusters 1 and 3 but not those of clusters 2 and 4, while 23S rRNA genes as targets covered all frankiae on the genus level, i.e., clusters 1, 2, 3, and 4 together. Targeting the 23S rRNA gene also allowed us to distinguish between cluster 1 and 3 frankiae and subgroups within cluster 1 (i.e., clusters 1a, 1b, and 1c) (15). The sum of individual detections generally equaled those on the genus level with both *nifH* and 23S rRNA genes as targets, indicating that members of clusters 2 and 4 were not present at all or not in detectable numbers in the soils analyzed (15, 17, 18). However, this statement is highly speculative, since direct proof of the presence or absence of cluster 2 and 4 frankiae in these soils has not been provided due to the lack of adequate detection and quantification procedures.

In this study, we have amended our database of sequences of 23S rRNA gene fragments of *Frankia* strains representing clusters 1 and 3 (15) to include sequences of frankiae assigned to clusters 2 and 4. Sequence alignments were used to design primers for the specific detection and quantification of *Frankia* clusters by Sybr Green-based qPCR and subsequently for the design of probes differentiating members of clusters 1, 2, and 3 (i.e., presumably all nitrogen-fixing frankiae) from those of cluster 4 (i.e., generally non-nitrogen-fixing frankiae). These probes were then used in combination with the specific primer combinations in TaqMan-based qPCR to quantify nitrogen-fixing (i.e., cluster 1, 2, and 3) and non-nitrogen-fixing (i.e., cluster 4) frankiae, as well as frankiae of the individual clusters and subgroups within cluster 1 in soils from different locations.

RESULTS

Comparative sequence analysis. Phylogenetic analysis of trimmed and aligned sequences provided a topology with four major clades that represented the previous assignment of clusters 1 to 4 (Fig. 1). Sequences of the uncultured endophytes from *Datisca*, *Coriaria*, and *Ceanothus* clustered with sequences of cluster 2 frankiae retrieved from the database, i.e., the endophyte from *Datisca glomerata* (Dg1) and pure-culture BMG5.1, representing one clade as cluster 2, while those of the atypical strains assembled together with cluster 4 strain EuI1c in another clade as cluster 4 (Fig. 1). While cluster 2 frankiae resembled a concise clade with uncorrected p distance values of aligned sequences between 94 and 100%, cluster 4 frankiae were more diverse, with 3 concise subgroups and distance values of aligned sequences between subgroups of 73 to 82% (Fig. 1).

Primer and probe design and evaluation. The amended database of partial 23S rRNA gene sequences of *Frankia* strains was used to design two forward primers, 23Dat1578f, targeting cluster 2 frankiae, and 23NNF1561f, targeting cluster 4 strains. Primer 23Dat1578f was specific with no mismatches to sequences of all cluster 2 frankiae and 3 to 7 mismatches to those of other *Frankia* strains in our database.

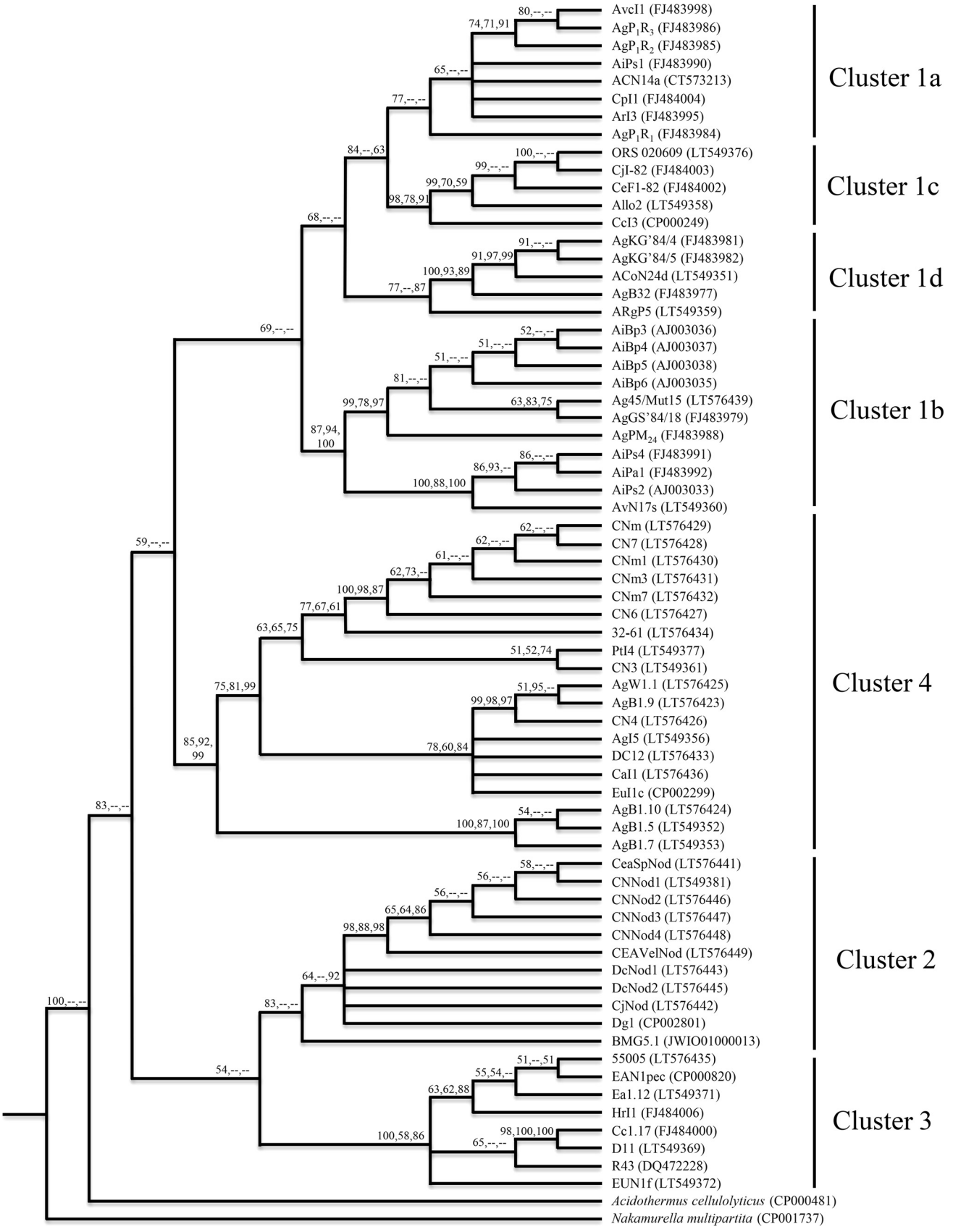

FIG 1 Sequence relationships for selected *Frankia* strains provided to show probe utility. Numbers above the branches represent the bootstrap values from a neighbor-joining (NJ) bootstrap analysis (10,000 replicates) using the HKY85 correction, followed by maximum likelihood bootstrap (1,000 replicates) and

(Continued on next page)

TestPrime 1.0 and TestProbe 3.0 analyses using the SILVA database revealed that primer 23Dat1578f (*Escherichia coli* position 1415) was specific for the two target organisms in the database, i.e., the *Frankia* endophyte from *Datisca glomerata* (Dg1; GenBank accession number CP002801) and strain BMG5.1 (JWIO01000013), and displayed at least 3 mismatches to sequences of nontarget frankiae (i.e., 3 mismatches for representatives of clusters 1c, CcI3, BMG5.23, and Thr and 4 mismatches for cluster 1a and 3 *Frankia* strains), while other organisms (*Prochlorococcus* sp.) displayed at least 5 mismatches. Together with primer 23Fra1769r, specific detection was achieved for the target organisms, i.e., cluster 2 frankiae, represented by the endophyte from *Datisca glomerata* (Dg1) in the database, with at least 4 mismatches to nontarget bacteria.

Primer 23NNF1561f (*E. coli* position 1415) could not be designed to cover all sequences of strains within cluster 4. It was specific for the subgroup including strains CN4, AgB1.9, AgW1.1, EuI1c, DC12, and CaI1, with 2 mismatches to strains CN3, CN6, CN7, CNm, CNm3, CNm7, 32-61, and PtI4 and 5 mismatches to AgB1.5, AgB1.7, and AgB1.10. TestPrime 1.0 and TestProbe 3.0 analyses of primer 23NNF1561f in the SILVA database only retrieved the sequence of strain EuI1c (CP002299) with no mismatches and confirmed 2 mismatches to that of strain CN3. For the primer target position on the 23S rRNA gene, identified as position 1415 on the reference gene of *E. coli*, nontarget sequences remained undetected at the program search limit of 5 mismatches. However, sequences with 3 and 4 mismatches to the sequences of nontarget organisms (e.g., *E. coli* and *Francisella tularensis*) were retrieved at a different position, i.e., *E. coli* position 419.

Probe NF1715f was designed to target all *Frankia* strains of clusters 1, 2, and 3 in our database; however, strains of cluster 3 (e.g., EUN1f, BMG5.12, and EAN1pec) displayed one mismatch. TestProbe analyses confirmed these data, with nontarget organisms (e.g., *Streptomyces* sp.) having at least 2 mismatches to probe NF1715f (*E. coli* position 1453). Probe NNF1715f (*E. coli* position 1462), targeting cluster 4 frankiae, was identical to sequences of cluster 4 strains detected by 23NNF1561f; however, they displayed 1 to 3 mismatches to strains of the remaining subgroups. TestProbe analyses retrieved only sequences from strain EuI1c without mismatches, while those of strain CN3 displayed 2 mismatches, while other nontarget organisms, such as *Actinoplanes* and *Streptomyces* sp., displayed 3 mismatches.

Annealing temperatures for all new primer combinations were established in the same range as those of our previously designed primers, 62 to 68°C (Table 1). In order to enhance coverage of primer 23NNF1561f to include members of the subgroup depicting 2 mismatches, annealing temperatures of 62°C were used instead of 66°C. Comparative analyses of amplifications using genus-specific detection with primers 1655f/1769r and specific detection with 23NNF1561f/1769r using pure cultures of AgB1.9, CN3, and AgB1.10, representing the three subgroups within cluster 4 and representative frankiae of the remaining clusters, resulted in complete detection of strains AgB1.9 and CN3, while strain AgB1.10 and other frankiae were not detected by the specific primer combination. The use of TaqMan-based qPCR using genus-specific primers 1655f/1769r and probes NF1715f and NNF1715f allowed us to circumvent coverage problems for cluster 4, since strains AgB1.9, CN3, and AgB1.10, representing the three subgroups, were detected quantitatively, as were strains representing nitrogen-fixing frankiae of the remaining subgroups.

Method evaluation. Amplicons of 23S rRNA genes of strains ArI3, Ag45/Mut15, CcI3, EAN1pec, AgB1.9, and uncultured frankiae from root nodules of *Coriaria* were specifically detected with primer combinations targeting the respective clusters (Fig. 2). Quantification of individual amplicons resulted in values comparable to those in mixtures of amplicons, independent of primers (i.e., genus or cluster specific) and

FIG 1 Legend (Continued)

Bayesian analysis values, respectively, noted for clades with greater than 50% bootstrap support. All results are plotted on the NJ bootstrap topology. Both the ML and Bayes analyses resolved generally similar topologies, although there are fewer supported tip nodes in the latter two analyses and the larger clades' basal relationships are either not resolved (ML) or have an alternative arrangement of the main clades (Bayes basal arrangement not depicted).

TABLE 1 Primer combinations for Sybr Green-based quantification of subgroups within the genus *Frankia*

Target group	Primer combination (5′→3′)	Annealing temp (°C)	Fragment size (bp)	Reference or source
Nitrogen-fixing *Frankia* strains of clusters 1 and 3; target gene *nifH*	nifHf1 (GGC AAG TCC ACC ACC CAG C) nifHr158 (GAC GCA CTT GAT GCC CCA)	64	191	16
Genus *Frankia* (clusters 1, 2, 3, and 4); target gene 23 rRNA	23Fra1655f (CTG GTA GTA GGC AAG CGA TGG) 23Fra1769r (GGC TCG GCA TCA GGT CTC AG)	64	133	15
Cluster 1 (*Alnus* and *Casuarina* host infection group)				
Subgroup 1a/d	23Ar1607f (GTG TCT TTT CGG AGA TGT GTC T) 23Fra1769r (GGC TCG GCA TCA GGT CTC AG)	64	128	17
Subgroup 1b	23Mut1555f (TTG ATG CGT CCA TGC TGA GG) 23Fra1769r (GGC TCG GCA TCA GGT CTC AG)	66	170	15
Subgroup 1c	23Cas1600f (GTG TCT CTT CGG AGG TGT GTT C) 23Fra1769r (GGC TCG GCA TCA GGT CTC AG)	68	128	17
Cluster 2 (*Rosaceae*/*Coriariaceae*/*Datiscaceae* host infection group)	23Dat1578f (TGG TTC GTG CTA ACC GTC CGA) 23Fra1769r (GGC TCG GCA TCA GGT CTC AG)	66	153	This study
Cluster 3 (*Elaeagnaceae*/*Rhamnaceae* host infection group)	23EAN1577f (GTT TGT GCT AAC CGT TCT GGT) 23Fra1769r (GGC TCG GCA TCA GGT CTC AG)	64	146	15
Cluster 4 (atypical, generally non-nitrogen-fixing and/or nonnodulating frankiae)	23NNF1561f (CCA ATG CTG AAT CTT CCT G) 23Fra1769r (GGC TCG GCA TCA GGT CTC AG)	62	142	This study

detection procedure (i.e., Sybr Green- or TaqMan-based qPCR) (Fig. 2). One-way analyses of variance (ANOVAs) did not detect statistically significant differences for cluster 1a ($P = 0.07$), 1b ($P = 0.3$), 1c ($P = 0.9$), 2 ($P = 0.1$), 3 ($P = 0.9$), and 4 ($P = 0.9$) across treatments. The sum of these clusters was also not statistically significantly different from genus- or group-specific detections ($P = 0.1$).

Sybr Green-based quantification using the *nifH* gene as the target to quantify frankiae of clusters 1 and 3 resulted in abundance estimates of 10×10^5 to 15×10^5 cells (g soil)$^{-1}$ depending on the soil (Fig. 3). Except for soil BAHF ($P = 0.002$), these estimates were not significantly different from those from 23S rRNA gene-based

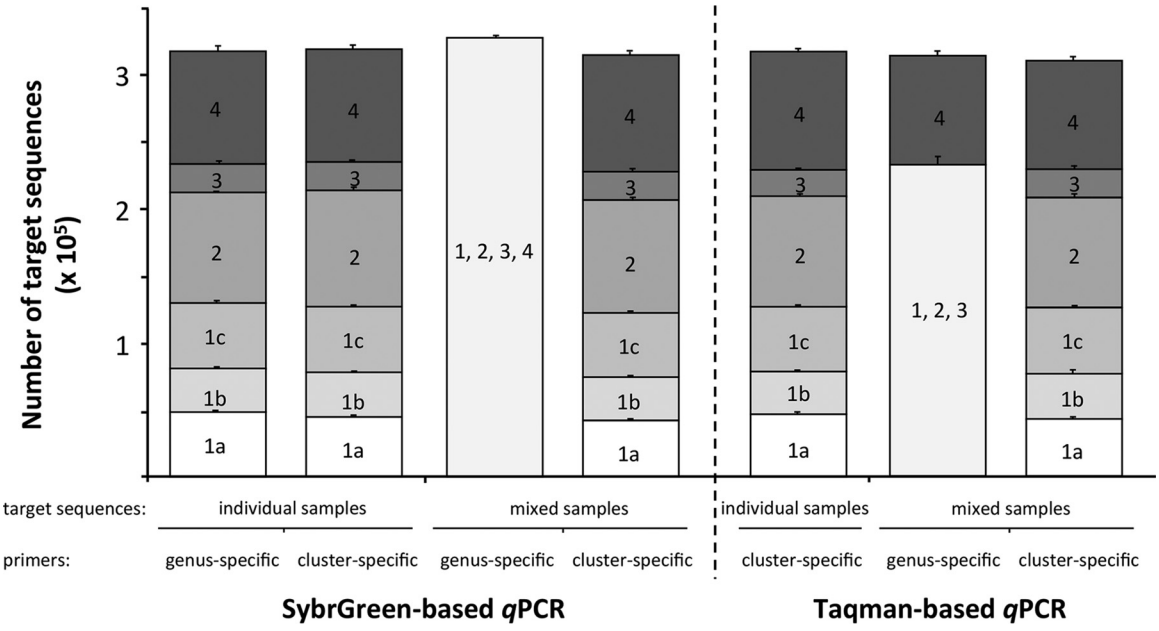

FIG 2 Sybr Green- and TaqMan-based qPCR quantification of amplicons from representative pure cultures or endophytes of *Frankia* clusters 1, 2, 3, and 4 and of subgroups a, b, and c within cluster 1. Individual amplicons or mixtures of amplicons were quantified using either genus- or (sub)cluster-specific primer combinations in Sybr Green-based analyses or additional probes targeting either cluster 1, 2, and 3 frankiae or cluster 4 frankiae.

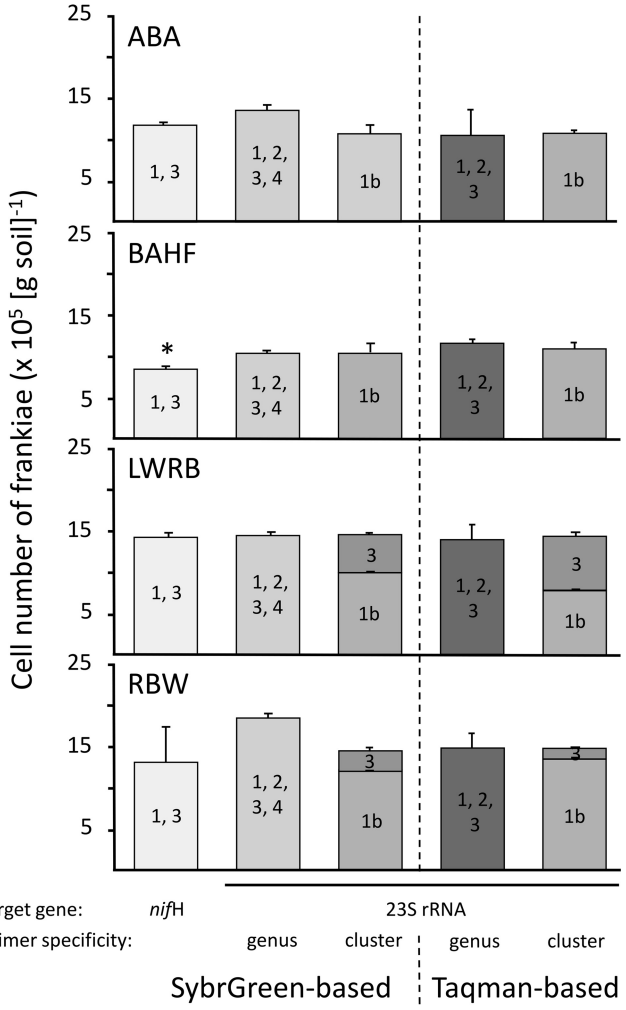

FIG 3 Sybr Green- and TaqMan-based qPCR quantification of *Frankia* clusters 1, 2, 3, and 4 in soils vegetated with host trees (*Alnus glutinosa*; soils ABA and BAHF) or nonhost trees (*Betula nigra*; soils LWRB and RBW). Quantification (from left to right) used *nifH* gene fragments as a target detecting clusters 1 and 3 or 23S rRNA gene fragments generated with primer combinations detecting the genus *Frankia*, i.e., all clusters 1, 2, 3, and 4 or primer combinations specific for clusters 1a/d, 1b, 1c, 2, 3, and 4 (presented as the sum of the individual clusters and subgroups detected). Only frankiae of subgroups 1b and cluster 3 were detected in these soils, while the remaining clusters and subgroups remained undetected (all using Sybr Green). TaqMan-based detection of 23S rRNA fragments resulted in compositions of frankiae with respect to clusters and subgroups that were similar to those of Sybr Green-based detection, with their sum representing quantification values similar to those of frankiae detected with genus-specific primers and probe NF1715f targeting frankiae of clusters 1, 2, and 3. Statistically significantly different values between treatments of the same sample are highlighted with an asterisk.

detection targeting frankiae of all four clusters or the sum of abundances of specific clusters and subgroups (*P* values between 0.07 and 0.9). The latter analyses revealed the presence of cluster 1 frankiae, especially subgroup 1b, represented by *Frankia* strain Ag45/Mut15, only or in combination with small numbers of frankiae of cluster 3 (Fig. 3). Frankiae of clusters 2 and 4, as well as cluster 1 frankiae, represented by strain Arl3 (subgroup 1a) and Ccl3 (subgroup 1c), were not detected. Abundance and diversity data obtained by Sybr Green-based analyses were confirmed in all four soils by TaqMan-based analyses of the individual clusters and subgroups, as well as by targeting all nitrogen-fixing frankiae, i.e., clusters 1, 2, and 3 (Fig. 3).

For prairie soils, Sybr Green-based detection using *nifH* as the target resulted in significantly lower abundance estimates (i.e., 5×10^5 to 10×10^5 cells [g soil]$^{-1}$) than Sybr Green-based detection of all clusters in the genus (i.e., 15×10^5 to 25×10^5 cells [g soil]$^{-1}$) ($P < 0.001$) (Fig. 4). Frankiae of clusters 1, 2, 3, and 4 were detected in two

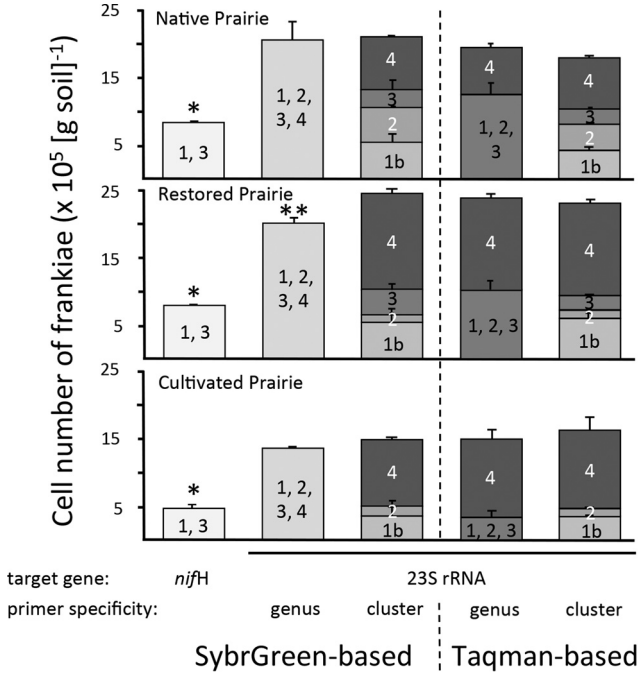

FIG 4 Sybr Green- and TaqMan-based qPCR quantification of *Frankia* clusters 1, 2, 3, and 4 in prairie soils (native, restored, and cultivated). Quantification (from left to right) used *nifH* gene fragments as a target detecting clusters 1 and 3 or 23S rRNA gene fragments generated with primer combinations detecting the genus *Frankia*, i.e., all clusters 1, 2, 3, and 4 or primer combinations specific for clusters 1a/d, 1b, 1c, 2, 3, and 4 (presented as the sum of the individual clusters and subgroups detected). Frankiae of subgroup 1b and clusters 2, 3, and 4 were generally detected in these soils, while subgroups 1a/d and 1c remained undetected (all using Sybr Green). Statistically significantly different values between treatments of the same sample are highlighted with one or two asterisks.

of the three soils, with the sum of their abundance estimates matching the estimates on the genus level (Fig. 4). TaqMan-based qPCR resulted in similar composition of the subgroups and detection of nitrogen-fixing frankiae equal to the sum of subgroups in clusters 1, 2, and 3 (Fig. 4). Subgroup 1a, represented by strain Arl3, and subgroup 1c, represented by strain Ccl3, were not detected at all and thus were not found to be abundant in any of the soils tested. Despite incomplete coverage of cluster 4 *Frankia* strains by both forward primer 23NNF1561f and probe NNF1715f, comparable abundance estimates for cluster 4 frankiae in these soils were obtained by both Sybr Green- and TaqMan-based qPCR (*P* values between 0.05 and 0.5) (Fig. 4).

DISCUSSION

Despite the short length of the sequences and thus low or no bootstrap support for major clades, the topology of our phylogenetic analyses confirmed cluster assignments of *Frankia* strains and uncultured endophytes from root nodules derived from comparative analyses of 16S rRNA gene sequences (4). Cluster assignments were also consistent with previous analyses using comparative sequence analyses of the insertion in domain III of the 23S rRNA gene (19), partial gyrase B (*gyrB*), nitrogenase reductase (*nifH*), or glutamine synthetase II (*glnII*) sequence analyses (20), as well as sequence analysis of the 16S-23S rRNA internally transcribed spacer (ITS) (21) or protein mass fingerprints from whole cells (22). While our analyses of uncultured endophytes in root nodules of *Datisca cannabina*, *Coriaria nepalensis*, *C. japonica*, and *Ceanothus* sp. revealed a limited sequence diversity and thus resulted in the assembly of a concise cluster 2, sequences of cluster 4 frankiae were more diverse, with significant sequence diversity establishing 3 subgroups.

Subgroups were previously described for clusters 1 and 3 (23, 24), with those of cluster 1 depicting remarkable physiological differences between each other (25).

Physiological characteristics of subgroups of cluster 3 remain to be studied. Within cluster 1, at least three subgroups were established (15), i.e., subgroup 1a, represented by *Frankia* strain Arl3, subgroup 1b, represented by *Frankia* strain Ag45/Mut15 that, in contrast to frankiae of subgroup 1a, was able to grow with leaf litter as the nutrient resource (25), and subgroup 1c, the *Casuarina*-infective strains that require the presence of host plants for growth (26). Our phylogenetic analysis retrieved an additional subgroup 1d (Fig. 1) that, however, could not be distinguished from subgroup 1a with the primer combinations used in our study (Table 1). Using sequences of the insertion in domain III of the 23S rRNA gene as the target for specific detection and quantification of these subgroups, subgroup 1b was found to be most prominent in many soils from temperate regions (15, 17, 18).

The lack of sequence diversity in our cluster 2 frankiae might be a function of our limited sampling strategy, i.e., the focus on uncultured *Frankia* endophytes in root nodules of a few host plants, i.e., *Datisca*, *Coriaria*, and *Ceanothus* sp. from locations in the United States and in Pakistan only. While previous studies on the diversity of cluster 2 frankiae in root nodules of different host plants also indicated low sequence diversity of these endophytes (27, 28), future studies on cluster 2 frankiae should include a larger diversity of host plants and more locations to retrieve additional information on *Frankia* diversity. Sequences of cluster 4 frankiae were more diverse, with 3 subgroups delineated. Since isolates are available for each of these subgroups, additional studies on physiological properties could provide information on a potential linkage between phenotypic and genotypic characteristics. Ultimately, however, additional assessments on overall diversity of the genus *Frankia* using next-generation sequencing methods in soil samples with and without host plants for all clusters will be required for a more comprehensive analysis of *Frankia* diversity.

Forward primers were designed identical to sequences of all cluster 2 frankiae (23Dat1578f) and those of one subgroup of cluster 4, represented by *Frankia* strain EuI1c (23NNF1561f). The latter subgroup was selected over the remaining 2 subgroups because isolates had been obtained from a variety of different host plant species, i.e., *Coriaria nepalensis*, *Alnus glutinosa*, *Elaeagnus umbellata*, *Datisca cannabina*, and *Ceanothus americanus* (CN4, AgB1.9, AgW1.1, EuI1c, DC12, and CaI1) (12, 13, 29–31). The reduction in annealing temperature from 66 to 62°C allowed us to increase coverage to two of the three subgroups detected without losing specificity for other *Frankia* strains. TaqMan-based quantification using genus-specific primers and a probe targeting cluster 4 frankiae resulted in the detection and reliable quantification of representative strains of these subgroups and thus might provide an adequate alternative to detection using specific forward primers. The new forward primers 23Dat1578f and 23NNF1561f, as well as the design of probes targeting nitrogen-fixing frankiae (clusters 1, 2, and 3) and cluster 4 frankiae (atypical, generally non-nitrogen-fixing strains), now allows us to expand qPCR analyses of frankiae in soils from clusters 1 and 3 and subgroups within cluster 1 to include clusters 2 and 4 and to distinguish between nitrogen-fixing and non-nitrogen-fixing populations (Table 1).

Previous analyses of soils ABA, BAHF, LWRB, and RBW revealed cell densities of about 10^6 cells (g soil)$^{-1}$, with cluster 1b representing the most prominent *Frankia* population, while cluster 3 frankiae were present in small numbers and clusters 1a, 1d, and 1c were usually absent (17). These results were largely confirmed in our current analyses, where specific analyses retrieved frankiae of cluster 1b only (soils ABA and BAHF), while soils LWRB and RBW harbored small numbers of cluster 3 frankiae as well (Fig. 3). Frankiae of clusters 2 and 4 remained undetected, indicating that these populations are either absent or present in numbers below the detection limit. Cluster 1b frankiae have been detected as major populations in several studies, with absolute numbers depending on the sampling depth, physicochemical conditions, and vegetation (15, 17, 18). These results demonstrated differential effects of environmental conditions, including plant species, carbon resources, and matric potentials on the fate of specific *Frankia* strains in soil. These factors could affect subpopulations of indigenous frankiae of clusters 2 and 4 that were both detected in all three prairie soils but

not in soils under *A. glutinosa* or *B. nigra*. Their presence in prairie soils but not in soils under host plant *A. glutinosa* and nonhost plant *B. nigra* suggests that vegetation is affecting the abundance of frankiae of these clusters. Populations of cluster 2 frankiae are lower in managed prairie soils, i.e., in restored and even more in cultivated prairie soils, compared to native prairie soils. Since native prairie harbors potential host plants for cluster 2 frankiae, such as *Ceanothus* species, it is tempting to assume that cluster 2 frankiae rely on the presence of host plants for growth. This situation would be similar to *Casuarina*-infective frankiae that, however, can persist long term in the absence of host plants. Long-term persistence in soil in the absence of host plants of cluster 2 frankiae would be in line with their detection in Tunisian soils lacking compatible host plants for more than 2 centuries, using bioassays with *Coriaria myrtifolia* as the capture plant (32). This assumption, however, needs to be assessed under controlled conditions, with different plant species as variables for population studies of cluster 2 frankiae.

Cluster 4 frankiae, i.e., presumably non-nitrogen-fixing frankiae, represented the most prominent *Frankia* population in all three prairie soils, while they were not detected at all in soils under *A. glutinosa* or *B. nigra*. Studies on cluster 4 frankiae are scarce, even though they have been reported to form an important fraction of all frankiae in wet soils under *A. glutinosa* (33, 34), with natural resistance to infection exhibited by different progenies of *A. glutinosa* (12, 35). While these results were based on bioassays, we now also have the instruments to assess the importance of cluster 4 frankiae in different environments or to use different strains of the non-nitrogen-fixing frankiae in controlled inoculation studies to retrieve information on their ecology in soils. Our initial results indicate that vegetation affects the basic composition of frankiae in soils, with higher diversity in prairie soils compared to much more restricted diversity under host and nonhost trees.

These results demonstrate the usefulness of the methodology developed, with the new forward primers and the probes providing more accurate coverage of the *Frankia* community in soils, even though members of cluster 2 and 4 might not be present in significant amounts in all soils. Future studies, however, should include next-generation sequencing analyses that assess the overall diversity of *Frankia* in a sequence of prairie soils, from native to highly managed, in order to determine environmental effects on diversity. In addition, our methodology now enables us to perform competition experiments that focus on the relationship between abundance and the nodule-forming capacity of cluster 1a/d and 1b frankiae, with frankiae of cluster 1a/d often found in nodules and cluster 1b frankiae dominant in soils.

MATERIALS AND METHODS

Cell sample preparation. Sequences of cluster 2 frankiae were obtained from uncultured endophytes of ethanol-preserved root nodules of *Datisca cannabina* collected in Rawalakot, Azad Kashmir, Pakistan (33.8472389, 73.7485194), *Coriaria nepalensis* was collected in Jhika Gali, Murree, District Rawalpindi, Pakistan (33.9112833, 73.4239306), *C. japonica* was from the Morton Arboretum in Lisle, IL, USA (41.8167861, −88.0679528), and *Ceanothus* sp. was from the Loda Cemetery Prairie Nature Preserve, IL, USA (40.5284721, −88.0717537). A single lobe was homogenized with a mortar and pestle in 1 ml of sterile water, and the homogenates were transferred to Eppendorf tubes and centrifuged at 14,000 × *g* for 1 min. The pellets were washed once with 0.1% (wt/vol) sodium pyrophosphate in water, followed by two washes with sterile distilled water. *Frankia* strains representing cluster 4 (CN3, CN4, CN6, CN7, CNm, CNm1, CNm3, CNm7, DC12, AgI5, AgW1.1, AgB1.5, AgB1.7, AgB1.9, and AgB1.10) (12, 13, 29, 36) were grown in P+N medium for 2 weeks (37), harvested by centrifugation at 14,000 × *g* for 5 min, and washed twice with sterile distilled water. Nodule pellets as well as pellets of pure cultures (approximately 50 mg) were resuspended in 95 μl of distilled water and lysed after addition of 5 μl of proteinase K solution (30 U mg^{-1}; 10 mg ml^{-1} in water; Promega, Madison, WI) and incubation at 37°C for 20 min (23). Afterwards, 0.5 μl of 10% SDS solution was added and the mixtures incubated at 37°C for another 3 h, which was followed by a final incubation at 80°C for 30 min (23).

PCR amplification. From these lysates, 2-μl aliquots were used as the template in subsequent PCR-based analyses. 23S rRNA gene fragments (about 240 bp) were amplified using primers 23Fra1533f (5′GTT GAT ATT CCC GTA CCG) and 23Fra1769r (5′GGC TCG GCA TCA GGT CTC AG), targeting frankiae and some other actinobacteria. The PCR was carried out in a volume of 50 μl, containing 1 μl of a 10 mM deoxynucleoside triphosphate (dNTP) mix, 0.5 μl each primer (0.4 μM), 8.2 μl bovine serum albumin (BSA) (30 μg ml^{-1}), 5 μl of 10 × PCR buffer with 15 mM $MgCl_2$, 2 μl root nodule or pure culture lysate, and 0.2 μl *Taq* DNA polymerase (5 U μl^{-1}; Gene Script, Piscataway, NJ) that was added after an initial incubation at 96°C for 10 min. The addition of *Taq* polymerase was followed by 35 rounds of temperature

cycling (96°C for 30 s, 60°C for 30 s, and 72°C for 45 s) and a final 7-min incubation at 72°C. Subsamples of the reaction mixtures (5 μl) were checked for amplification products by gel electrophoresis (1%, wt/vol, agarose in Tris-acetate-EDTA buffer) after staining with ethidium bromide (0.5 μg ml^{-1}) (38).

Sequence analyses. Amplified 23S rRNA gene fragments were cleaned using shrimp alkaline phosphatase and exonuclease I (Affymetrix, Santa Clara, CA) by following the manufacturer's protocols and then sequenced bidirectionally using BigDye Terminator v3.1 (Applied Biosystems, Foster City, CA) with the same primers used for PCR. Sequences were analyzed on a 3500 genetic analyzer (Life Technologies, Carlsbad, CA) and deposited at GenBank under accession numbers LT576423 to LT576449.

Phylogenetic analyses. Sequences of amplified 23S rRNA gene fragments obtained from uncultured frankiae from root nodules of all plants analyzed and those of pure cultures of cluster 4 *Frankia* were trimmed to lengths between 141 and 152 bp to match those in our database (19, 39), assembled in Geneious 9.1.4 (Biomatters Ltd., Auckland, New Zealand), and checked in GenBank/EMBL databases using the BLAST algorithm (40). Representative sequences from confirmed *Frankia* strains of all 4 clusters were added from our and GenBank/EMBL databases and aligned by the Geneious alignment tool. The identity and relationship among the sequences amplified were evaluated using neighbor-joining (NJ) (41), maximum likelihood (ML) (42), and Bayesian analyses (43). All of these analyses were conducted from within Geneious 9.1.4. The neighbor-joining analyses utilized the HKY85 model to correct for substitution bias (49). Model parameters for maximum likelihood, which were estimated by the general time-reversible model (GTR) with gamma (44), were used as input in an ML heuristic search using RAxML (45). Bootstrap values (46) were estimated from a heuristic search with random stepwise addition of sequence for 10,000 NJ and 1,000 ML iterations. MrBayes version 3.1.2 (43) was implemented for 10 million generations, saving every thousandth tree, with a burn-in of one million trees using the GTR with gamma substitution model.

Primer and probe design and evaluation. Aligned sequences were amended with sequences of other target and nontarget organisms and used to manually check for and design forward primers specific for cluster 2 and 4 frankiae, i.e., primers 23Dat1578f and 23NNF1561, respectively, which could be used with reverse primer 23Fra1769r in Sybr Green-based qPCR (Table 1). In addition, two probes, one targeting all frankiae of clusters 1, 2, and 3 (and thus supposedly all nitrogen-fixing frankiae), probe NF1715f (5′-6-carboxyfluorescein [FAM]-TGG TTG TCC TGG GGC AAG GGT GTA GG-6-carboxytetramethylrhodamine [TAMRA]), and a second targeting cluster 4 frankiae (and thus generally non-nitrogen-fixing frankiae), probe NNF1715f (5′-6-FAM-CGG GGT AAG CGT GTA GG ACG ACG TGT A-TAMRA), were designed and subsequently evaluated in TaqMan-based qPCR using the genus- or subgroup-specific primer sets from Sybr Green-based applications for amplification. Selected primers and probes were checked for low potential of self- and heterodimer formation using OligoAnalyzer 3.1 (www.idtdna.com/calc/analyzer) and for target specificity using TestPrime 1.0 and TestProbe 3.0 (47) from the SILVA rRNA database project (www.arb-silva.de; accessed 29 September 2016) (48). Annealing temperatures for all primer combinations were tested in qPCRs with DNA of representative *Frankia* strains or PCR products from uncultured endophytes of clusters 1, 2, 3, and 4, respectively, and quantifications were compared between Sybr Green- and TaqMan-based analyses, both performed in an Eco real-time PCR system (Illumina, San Diego, CA).

Primer combinations for Sybr Green-based qPCR targeted *nifH* gene sequences (16) or 23S rRNA gene sequences (15) (Table 1). Sybr Green-based analyses were carried out in triplicate in a total volume of 10 μl containing 5 μl of SsoADV Sybr Green mix (Bio-Rad, Hercules, CA), 0.125 μl of forward and reverse primers (100 nM each), and 1 μl of DNA template using an initial denaturation at 95°C for 5 min and 40 cycles of denaturation at 95°C, annealing at 62, 64, or 66°C depending on the primer combination (Table 1), and extension at 72°C, each for 30 s (15, 17). The amplification was followed by a melting curve analysis.

Primer combinations targeting 23S rRNA sequences were also used for TaqMan-based quantification, though in combination with probe NF1715f or NNF1715f. Except for cluster 1b, all TaqMan-based analyses were carried out in triplicate in a volume of 10 μl containing 5 μl of Sso ADV probe mix (Bio-Rad), 0.2 μl of forward and reverse primers (100 nM each), 0.25 μl of probe (250 nM each), and 1 μl of DNA template. An initial denaturation at 95°C for 5 min was followed by 40 cycles of 60°C for 60 s. For cluster 1b, primer concentrations were 300 nM each, and cycles consisted of 58°C for 60 s followed by 72°C for 30 s.

Quantification was based on standard curves generated from purified PCR products of *nifH* or 23S rRNA genes of strains Ag45/Mut15, Arl3, Ccl3, EAN1pec, and AgB1.9 or uncultured frankiae from root nodules of *Coriaria*, depending on the primer combination. Amplicons were generated using the genus-specific primers (Table 1), and concentrations were measured with a Qubit 2.0 fluorometer (Life Technologies, Carlsbad, CA). Copy numbers were calculated from concentrations (http://cels.uri.edu/gsc/cndna.html) and normalized after qPCR quantification with the primer combination targeting all nitrogen-fixing frankiae. Copy numbers were divided by copy numbers of the 23S rRNA gene per genome (2 or 3 for pure cultures, 2.5 for unknown populations) to relate copy numbers to *Frankia* cell numbers (15).

Method evaluation. To assess probe specificity and effects of detection procedures on the quantification of frankiae, amplicons of 23S rRNA genes of strains Ag45/Mut15, Arl3, Ccl3, EAN1pec, and AgB1.9 and uncultured frankiae from root nodules of *Coriaria* were generated using genus-specific primers. These amplicons were initially used as individual amplicons or in mixtures in Sybr Green-based qPCR to compare quantification with either genus- or cluster-specific primers. This approach was extended by the addition of probes in TaqMan-based qPCR.

Further method assessments used soil samples that were obtained from seven sites in Illinois, in close proximity to Urbana-Champaign. Soils included 4 previously analyzed sites, i.e., sites ABA (Arboretum at the University of Illinois; 40.093585, −88.218016) and BAHF (Horticulture Farm at the University of Illinois; 40.079306, −88.190558), planted with European alder (*Alnus glutinosa*), and sites LWRB (Lake of the Woods Park; 40.203501, −88.387924) and RBW (Illinois State Water Survey Campus; 40.083917, −88.242038), planted with river birch (*Betula nigra*) (17). Soils at sites ABA, BAHF, and RBW formed under tallgrass prairie on postglacial loess, while the soil at site LWRB formed under deciduous forest, all about 23,000 years before present. Additional soils were obtained from two sites at Loda Cemetery Prairie (40.5284721, −88.0717537), one representing native prairie dominated by *Sorghastrum nutans* (L.) (Indiangrass) and *Andropogon gerardii* Vitman (Big bluestem) on black prairie soil, while the second was adjacent to the native prairie site but cultivated continuously with corn. The last soil was obtained from Meadow Brook Park (40.0789008, −83.7852567) and resembled restored prairie dominated by *S. nutans* (L.) and *A. gerardii* Vitman on black prairie soil. All soils were similar with respect to particle size distribution (silt loam), organic matter content (2.5 to 4.6%), and pH (6.0 to 7.1). At each site, soil samples of about 1 kg were taken from the upper 10 cm, with soils from sites with trees (ABA, BAHF, LWRB, and RBW) being sampled less than 1 m from the stem of one tree. Samples were obtained using a trowel that was cleaned with a wire brush and then rinsed in a bucket containing 50% ethanol between sample extractions. Soils were released from roots and homogenized by manipulating the entire sample in freezer bags and then stored at 4°C until further processing.

DNA was extracted from triplicate 250-mg (dry weight) soil samples using the SurePrep soil DNA isolation kit (Fisher Scientific, Houston, TX) with small modifications as described before (16). Tenfold dilutions were used as the template in both Sybr Green- and TaqMan-based qPCR analyses for members of the genus *Frankia* or subgroups within the genus, as described above. Results of all analyses were corrected for extraction efficiencies determined as the ratio of inoculated *Salmonella enterica* serovar Typhimurium (ATCC 14028) cells detected by qPCR-based quantification of a 268-bp *invA* gene fragment before and after extraction as described previously (16).

Statistical analysis. One-way ANOVA and pairwise multiple-comparison procedures (Holm-Sidak method) were used in SigmaPlot 13.0 (Systat Software Inc., San Jose, CA) to assess the effects of different qPCR procedures on abundance estimates for frankiae, with a significance level at a P value of <0.05.

Accession number(s). Sequences determined here were deposited at GenBank under accession numbers LT576423 to LT576449.

ACKNOWLEDGMENTS

We are indebted to the Office of Sponsored Programs (Research Enhancement Program grant no. 9000000647), the Graduate College, and the Department of Biology at Texas State University for financial support.

REFERENCES

1. Schwintzer CR, Tjepkema JD. 1990. The biology of *Frankia* and actinorhizal plants. Academic Press, San Diego, CA.
2. Huss-Danell K. 1997. Actinorhizal symbioses and their N-2 fixation. New Phytol 136:375–405. https://doi.org/10.1046/j.1469-8137.1997.00755.x.
3. Benson DR, Silvester WB. 1993. Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. Microbiol Rev 57:293–319.
4. Normand P, Orso S, Cournoyer B, Jeannin P, Chapelon C, Dawson J, Evtushenko L, Misra AK. 1996. Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family Frankiaceae. Intl J System Bacteriol 46:1–9. https://doi.org/10.1099/00207713-46-1-1.
5. Benson DR, Dawson J. 2007. Recent advances in the biogeography and genecology of symbiotic *Frankia* and its host plants. Physiol Plant 130:318–330. https://doi.org/10.1111/j.1399-3054.2007.00934.x.
6. Dobritsa SV. 1998. Grouping of *Frankia* strains on the basis of susceptibility to antibiotics, pigment production and host specificity. Int J Syst Bacteriol 48:1265–1275. https://doi.org/10.1099/00207713-48-4-1265.
7. Callaham D, Del Tredici P, Torrey JG. 1978. Isolation and cultivation *in vitro* of the actinomycete causing root nodulation in *Comptonia*. Science 199:899–902. https://doi.org/10.1126/science.199.4331.899.
8. Diem HG, Gauthier D, Dommergues YR. 1982. Isolation of *Frankia* from nodules of *Casuarina equisetifolia*. Can J Microbiol 28:526–530. https://doi.org/10.1139/m82-079.
9. Weber A, Smolander A, Nurmiaho-Lassila EL, Sundman V. 1988. Isolation and characterization of *Frankia* strains from *Alnus incana* and *Alnus glutinosa* in Finland. Symbiosis 6:97–116.
10. Caru M, Becerra A, Sepulveda D, Cabello A. 2000. Isolation of infective and effective *Frankia* strains from root nodules of *Alnus acuminata* (Betulaceae). World J Microbiol Biotechnol 16:647–651. https://doi.org/10.1023/A:1008929520571.
11. Gtari M, Brusetti L, Skander G, Mora D, Boudabous A, Daffonchio D. 2004. Isolation of *Elaeagnus*-compatible *Frankia* from soils collected in Tunisia. FEMS Microbiol Lett 234:349–355. https://doi.org/10.1111/j.1574-6968.2004.tb09554.x.
12. Hahn D, Starrenburg MJC, Akkermans ADL. 1988. Variable compatibility of cloned *Alnus glutinosa* ecotypes against ineffective *Frankia* strains. Plant Soil 107:233–243. https://doi.org/10.1007/BF02370552.
13. Mirza MS, Hahn D, Akkermans ADL. 1992. Isolation and characterization of *Frankia* strains from *Coriaria nepalensis*. Syst Appl Microbiol 15:289–295. https://doi.org/10.1016/S0723-2020(11)80103-7.
14. Gtari M, Ghodhbane-Gtari F, Nouioui I, Ktari A, Hezbri K, Mimouni W, Sbissi I, Ayari A, Yamanaka T, Normand P, Tisa LS, Boudabous A. 2015. Cultivating the uncultured: growing the recalcitrant cluster-2 *Frankia* strains. Sci Rep 5:13112. https://doi.org/10.1038/srep13112.
15. Samant S, Amann RI, Hahn D. 2014. Evaluation of the 23S rRNA gene as target for qPCR based quantification of *Frankia* in soils. System Appl Microbiol 37:229–234. https://doi.org/10.1016/j.syapm.2013.11.001.
16. Samant S, Sha Q, Iyer A, Dhabekar P, Hahn D. 2012. Quantification of *Frankia* in soils using SYBR Green based qPCR. Syst Appl Microbiol 35:191–197. https://doi.org/10.1016/j.syapm.2011.12.004.
17. Samant S, Huo T, Dawson JO, Hahn D. 2016. Abundance and relative distribution of *Frankia* host infection groups under actinorhizal *Alnus glutinosa* and non-actinorhizal *Betula nigra* trees. Microb Ecol 71:473–481. https://doi.org/10.1007/s00248-015-0643-2.
18. Samant SS, Dawson JO, Hahn D. 2015. Growth responses of indigenous *Frankia* populations to edaphic factors in actinorhizal rhizospheres. Syst Appl Microbiol 38:501–505. https://doi.org/10.1016/j.syapm.2015.07.005.
19. Hönerlage W, Hahn D, Zepp K, Zeyer J, Normand P. 1994. A hypervariable 23S rRNA region provides a discriminating target for specific characterization of uncultured and cultured *Frankia*. Syst Appl Microbiol 17:433–443. https://doi.org/10.1016/S0723-2020(11)80061-5.
20. Nouioui I, Ghodhbane-Gtari F, Beauchemin NJ, Tisa LS, Gtari M. 2011.

Phylogeny of members of the *Frankia* genus based on *gyr*B, *nif*H and glnII sequences. Antonie Van Leeuwenhoek 100:579–587. https://doi.org/10.1007/s10482-011-9613-y.
21. Ghodhbane-Gtari F, Nouioui I, Chair M, Boudabous A, Gtari M. 2010. 16S-23S rRNA intergenic spacer region variability in the genus *Frankia*. Microb Ecol 60:487–495. https://doi.org/10.1007/s00248-010-9641-6.
22. Hahn D, Mirza B, Benagli C, Vogel G, Tonolla M. 2011. Typing of nitrogen-fixing *Frankia* strains by matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry. Syst Appl Microbiol 34:63–68. https://doi.org/10.1016/j.syapm.2010.11.009.
23. Mirza BS, Welsh A, Rasul G, Rieder JP, Paschke MW, Hahn D. 2009. Variation in *Frankia* populations of the *Elaeagnus* host infection group in nodules of six host plant species after inoculation with soil. Microb Ecol 58:384–393. https://doi.org/10.1007/s00248-009-9513-0.
24. Welsh A, Mirza BS, Rieder JP, Paschke MW, Hahn D. 2009. Diversity of frankiae in root nodules of *Morella pensylvanica* grown in soils from five continents. Syst Appl Microbiol 32:201–210. https://doi.org/10.1016/j.syapm.2009.01.002.
25. Mirza BS, Welsh AK, Hahn D. 2009. Growth of *Frankia* strains in leaf litter-amended soil and the rhizosphere of a non-actinorhizal plant. FEMS Microbiol Ecol 70:132–141. https://doi.org/10.1111/j.1574-6941.2009.00746.x.
26. Samant S, Dawson JO, Hahn D. 2016. Growth responses of introduced *Frankia* strains to edaphic factors. Plant Soil 400:123–132. https://doi.org/10.1007/s11104-015-2720-1.
27. Vanden Heuvel BD, Benson DR, Bortiri E, Potter D. 2004. Low genetic diversity among *Frankia* spp. strains nodulating sympatric populations of actinorhizal species of *Rosaceae*, *Ceanothus* (*Rhamnaceae*) and *Datisca glomerata* (*Datiscaceae*) west of the Sierra Nevada (California). Can J Microbiol 50:989–1000. https://doi.org/10.1139/w04-079.
28. Oakley B, North M, Franklin JF, Hedlund BP, Staley JT. 2004. Diversity and distribution of *Frankia* strains symbiotic with *Ceanothus* in California. Appl Environ Microbiol 70:6444–6452. https://doi.org/10.1128/AEM.70.11.6444-6452.2004.
29. Mirza MS, Janse JD, Hahn D, Akkermans ADL. 1991. Identification of atypical *Frankia* strains by fatty acid analysis. FEMS Microbiol Lett 83:91–97. https://doi.org/10.1111/j.1574-6968.1991.tb04395.x.
30. Lechevalier MP, Ruan JS. 1984. Physiology and chemical diversity of *Frankia* spp. isolated from nodules of *Comptonia peregrina* (L.) Coult. and *Ceanothus americanus* L. Plant Soil 78:15–22. https://doi.org/10.1007/BF02277836.
31. Baker D, Newcomb W, Torrey JG. 1980. Characterization of an ineffective actinorhizal microsymbiont, Frankia sp Eul1 (Actinomycetales). Can J Microbiol 26:1072–1089. https://doi.org/10.1139/m80-180.
32. Nouioui I, Sbissi I, Ghodhbane-Gtari F, Benbrahim KF, Normand P, Gtari M. 2013. First report on the occurrence of the uncultivated cluster 2 *Frankia* microsymbionts in soil outside the native actinorhizal host range area. J Biosci 38:695–698. https://doi.org/10.1007/s12038-013-9366-z.
33. Wolters DJ, Akkermans ADL, Van Dijk C. 1997. Ineffective *Frankia* strains in wet stands of *Alnus glutinosa* L Gaertn in the Netherlands. Soil Biol Biochem 29:1707–1712. https://doi.org/10.1016/S0038-0717(97)00075-8.
34. Wolters DJ, Van Dijk C, Zoetendal EG, Akkermans AD. 1997. Phylogenetic characterization of ineffective *Frankia* in *Alnus glutinosa* (L.) Gaertn. nodules from wetland soil inoculants. Mol Ecol 6:971–981. https://doi.org/10.1046/j.1365-294X.1997.00265.x.
35. Van Dijk C, Sluimer A. 1994. Resistance to an ineffective *Frankia* strain type in *Alnus glutinosa* (L.) Gaertn. New Phytol 128:497–504. https://doi.org/10.1111/j.1469-8137.1994.tb02996.x.
36. Mirza MS, Akkermans WM, Akkermans ADL. 1994. PCR-amplified 16S rRNA sequence analysis to confirm nodulation of *Datisca cannabina* L. by the endophyte of *Coriaria nepalensis* Wall. Plant Soil 160:147–152. https://doi.org/10.1007/BF00150356.
37. Meesters TM, van Genesen ST, Akkermans ADL. 1985. Growth, acetylene reduction activity and localization of nitrogenase in relation to vesicle formation in *Frankia* strains Cc1.17 and Cp1.2. Arch Microbiol 143:137–142. https://doi.org/10.1007/BF00411036.
38. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Maunuksela L, Zepp K, Koivula T, Zeyer J, Haahtela K, Hahn D. 1999. Analysis of *Frankia* populations in three soils devoid of actinorhizal plants. FEMS Microbiol Ecol 28:11–21. https://doi.org/10.1111/j.1574-6941.1999.tb00556.x.
40. Pearson WR, Lipman DJ. 1988. Improved tools for biological sequence comparison. Proc Natl Acad Sci U S A 85:2444–2448. https://doi.org/10.1073/pnas.85.8.2444.
41. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.
42. Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690. https://doi.org/10.1093/bioinformatics/btl446.
43. Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574. https://doi.org/10.1093/bioinformatics/btg180.
44. Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Am Math Soc Lect Math Life Sci 17:57–86.
45. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
46. Felsenstein J. 1985. Confidence limits of phylogenies: an approach using the bootstrap. Evolution 39:783–791. https://doi.org/10.2307/2408678.
47. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41:e1. https://doi.org/10.1093/nar/gks808.
48. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner F-O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596. https://doi.org/10.1093/nar/gks1219.
49. Hasegawa M, Kishino H, Yano T. 1985. Dating of human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174.